

Assignment games with externalities revisited

Jens Gudmundsson¹  · Helga Habis²

Received: 2 December 2016 / Accepted: 17 March 2017
© The Author(s) 2017. This article is an open access publication

Abstract We study assignment games with externalities. The value that a firm and a worker create depends on the matching of the other firms and workers. We ask how the classical results on assignment games are affected by the presence of externalities. The answer is that they change dramatically. Though stable outcomes exist if agents are “pessimistic”, this is a knife-edge result: we show that there are problems in which the slightest optimism by a single pair erases all stable outcomes. If agents are sufficiently optimistic, then there need not exist stable outcomes even if externalities are vanishingly small. The negative result persists also when we impose a very restrictive structure on the values and the externalities. Furthermore, stability and efficiency no longer go hand in hand and the set of stable outcomes need not form a lattice with respect to the agents’ payoffs.

Keywords Two-sided matching · Assignment games · Externalities · Stability

We thank Tommy Andersson, Federico Echenique, Bettina Klaus, and Flip Klijn as well as participants at the Workshop on Networks and Externalities, the Summer School on Matching Problems in Budapest, 2013, the UECE Lisbon Meetings, 2013, and at Social Choice and Welfare, 2014, for valuable comments. Gudmundsson gratefully acknowledges the financial support of the Jan Wallander and Tom Hedelius Foundation. Habis would like to thank the financial support of OTKA-112266 through Corvinus University of Budapest.

✉ Jens Gudmundsson
jens.gudmundsson@nek.lu.se

Helga Habis
helga.habis@uni-corvinus.hu

¹ Department of Economics, Lund University, Box 7082, 222 07 Lund, Sweden

² Department of Microeconomics, Corvinus University of Budapest, Fővám tér 8., Budapest 1093, Hungary

1 Introduction

We examine the impact that externalities have on assignment games (Koopmans and Beckmann 1957; Shapley and Shubik 1971). Assignment games provide a primitive model of the job market in which firms hire workers after negotiating salaries; we refer to Roth and Sotomayor (1992) for a survey. Externalities are modelled by conditioning the success of a firm and a worker on the matching of the other firms and workers. So, how does this generalization affect the classical results on assignment games? Our results suggest—dramatically. In particular, all of the findings of Shapley and Shubik (1971) are overturned. In summary, (i) a problem may have no stable outcome, (ii) efficient outcomes can be unstable, and (iii) the set of stable outcomes need not form a lattice with respect to the agents' payoffs.

We are motivated by the prevalence of externalities in everyday practical examples. Consider a firm that produces phones. Its success stems of course mainly from the competence of its employees, but it may also be influenced by firms that produce complementary goods (like network capacity and signal) and its rivals that produce similar phones. Thus, the value that a firm and a worker create at a matching—that is, the amount of money that they divide into profit and salary—depends on how the other agents are matched.

In our model, an outcome is a matching, a profit for each firm, and a salary for each worker. We study outcomes that cannot be *blocked*, that is, *stable* outcomes. A firm f and a worker w block an outcome if they are better off together than with their current partners. Note, however, that blocking is not as straightforward as it may appear at first glance. Upsetting the current matching affects not only the blocking pair f and w , but also the values of the other pairs. As a consequence, other agents may swap partners and thereby affect the value available to f and w .¹ Hence, how agents forecast that other agents react to their block is crucial to what is stable.

To decide on whether to block, f and w take all contingencies into account simultaneously. At each matching that can be formed upon a block by f and w , f and w create some value. According to their attitude towards risk or beliefs about how the other agents will match, these values induce a *blocking threshold*. The pair may condition its threshold on the current matching, so it can, say, assign higher probability on matchings more “similar” to the current. One interpretation of the threshold is that it reflects the value that f and w expect to create if they block the current outcome. If the threshold exceeds the sum of their current payoffs, then they block.

We can distinguish different types of agents based on how they determine their threshold. A pair is *pessimistic* if the pair's threshold is the smallest value that the pair can create by matching. Thus, the pair is (maximally) pessimistic in its forecast of the other agents' reaction. Alternatively, the pair blocks only if it with certainty is made better off. If pairs always are pessimistic, then there exist stable outcomes (due to

¹ This is typically referred to as *residual behaviour* in the literature on cooperative game theory. See for instance Aumann and Peleg (1960) and Shenyoy (1980).

Sasaki and Toda 1996). We complement this by showing necessity of pessimism. Theorem 1 states that, if there is just one pair that at some matching is not pessimistic, then there are values such that the corresponding assignment game with externalities lacks stable outcomes.² This finding is strong as it requires only the smallest conceivable form of optimism but does so by exploiting that externalities may be arbitrarily large. The “opposite” strategy can be applied to show that, if agents are sufficiently optimistic, then there need not exist stable outcomes even if externalities are vanishingly small (Example 2). Furthermore, even if externalities are “reasonably” bounded, the negative result persists (Example 5). In contrast to games without externalities, stability and efficiency no longer go hand in hand: we provide a stable but inefficient matching and an efficient but unstable matching (Example 3). Another discrepancy compared to games without externalities is that the set of stable outcomes need not form a lattice with respect to the agents’ payoffs (Example 4).

Sasaki and Toda (1996) mainly study externalities in the marriage problem (Gale and Shapley 1962),³ though they also briefly examine assignment games focusing exclusively on pessimistic agents. Their approach to the marriage problem is to endow each agent i with an “estimation function” φ_i that to each potential partner j associates a set of matchings at which i and j are matched. Agents f and w form a blocking pair whenever they prefer each matching $\mu \in \varphi_f(w) \cup \varphi_w(f)$ to the current matching. This approach is very reasonable as, in the ordinal marriage problem, the average of two outcomes (that is, two matchings) bears no meaning. Therefore, blocking agents cannot compare the current matching with, say, the expected post-block matching. In contrast, the primitives of assignment games—the values—contain cardinal information. Therefore, it is possible and meaningful for the agents to take all contingencies into account simultaneously rather than separately, as agents do in our approach. Thus, not only is our approach more flexible and general than using estimation functions, but it is also better tailored to assignment games.

A series of recent papers are related to ours. Chen’s (2013) work is primarily focused on applications, though he also finds a negative result that parallels our Theorem 1. However, as his notion of stability uses estimation functions in the spirit of Sasaki and Toda (1996), his result is subsumed by ours. Pycia and Yenmez (2016) examine externalities in a matching with contracts-setting. They find that, if choice functions are substitutable and satisfy “irrelevance of rejected alternatives”, then stable outcomes (sets of contracts) exist. In contrast to our work, Pycia and Yenmez (2016) consider a notion of stability in line with Gale and Shapley (1962) in the sense that agents assume that their choices do not trigger reactions by the other agents. Instead, such predictions can be built into the choice functions. Last, Eriksson et al. (2011) model assignment games with social preferences. The value that a pair creates is independent of the matching, but agents’ utilities depend negatively on the payoffs of the other agents on the same side (for example, firms affect other firms). In our framework, payoffs and utilities coincide. Eriksson et al. (2011) show that several results carry over from

² Results of a similar nature have been found for cooperative games (Funaki and Yamato 1999; Kóczy 2007) and housing markets (Mumcu and Saglam 2007).

³ That is, two-sided matching without money. See also Li (1993), Dutta and Massó (1997), Echenique and Yenmez (2007), and Hafalir (2008).

Shapley and Shubik (1971) if agents are *naïve*. This is an assumption of bounded rationality that, in essence, implies that agents expect almost everything to stay the same after blocking. As just discussed in relation to Pycia and Yenmez (2016), this is in contrast to our approach.

The outline of the paper is as follows: we introduce the model in Sect. 2. In Sect. 3, we discuss the existence of stable outcomes, their relation with efficiency, and the structure of the set of stable outcomes. In Sect. 4, we impose more structure on the model but find that pessimism still is a necessity for the existence of stable outcomes.

2 Model

There are m firms F and $n = m$ workers W .⁴ A **matching** is a bijection $\mu : F \rightarrow W$. The **set of matchings** is \mathcal{M} . Let $\mathcal{M}_{fw} = \{\mu \in \mathcal{M} : \mu(f) = w\}$ and $\mathcal{M}_{-fw} = \mathcal{M} \setminus \mathcal{M}_{fw}$. If $\mu(f) = w$, then we say that $(f, w) \in \mu$. For each $(f, w) \in F \times W$, $\alpha_{fw} : \mathcal{M}_{fw} \rightarrow \mathbb{R}_+$ maps a **value** to each matching. Note that the value created by a pair may differ across matchings. Let $A \equiv (\alpha_{fw})_{f \in F, w \in W}$ and let \mathcal{A} be the collection of such A 's.

A **payoff vector** for F is $u \in \mathbb{R}_+^m$, where u_f is firm f 's profit. Likewise, a payoff vector for W is $v \in \mathbb{R}_+^n$, where v_w is worker w 's salary. Payoff vectors $(u, v) \in \mathbb{R}_+^m \times \mathbb{R}_+^n$ are **compatible** with $\mu \in \mathcal{M}$ in $A \in \mathcal{A}$ if, for each $f \in F$ and $w = \mu(f)$, $u_f + v_w = \alpha_{fw}(\mu)$. An **outcome** of $A \in \mathcal{A}$ is $(\mu, u, v) \in \mathcal{M} \times \mathbb{R}_+^m \times \mathbb{R}_+^n$ such that (u, v) are compatible with μ in A . Given $A \in \mathcal{A}$, $\mu \in \mathcal{M}$ is **efficient** if, for each $\mu' \in \mathcal{M}$,

$$\sum_{(f,w) \in \mu} \alpha_{fw}(\mu) \geq \sum_{(f,w) \in \mu'} \alpha_{fw}(\mu').$$

If, with some abuse of notation, for each $(f, w) \in F \times W$ and $\{\mu, \mu'\} \subseteq \mathcal{M}$, $\alpha_{fw}(\mu) = \alpha_{fw}(\mu') \equiv \alpha_{fw}$, then *there are no externalities*. The **collection of (assignment) games without externalities** is \mathcal{A}^0 . A pair $(f, w) \in F \times W$ **blocks** the outcome $(\mu, u, v) \in \mathcal{M} \times \mathbb{R}_+^m \times \mathbb{R}_+^n$ of $A \in \mathcal{A}^0$ if $u_f + v_w < \alpha_{fw}$. An outcome (μ, u, v) is **stable** in $A \in \mathcal{A}^0$ if no pair blocks it.

Proposition 1 (Shapley and Shubik 1971) *Let $A \in \mathcal{A}^0$ and $(\mu, u, v) \in \mathcal{M} \times \mathbb{R}_+^m \times \mathbb{R}_+^n$.*

1. *If (μ, u, v) is stable in A , then μ is efficient in A .*
2. *If (μ, u, v) is stable in A and $\mu' \in \mathcal{M}$ is efficient, then (μ', u, v) is stable in A .*
3. *The set of stable outcomes of A forms a non-empty complete lattice with respect to the firms' profits and the workers' salaries.*

We ask if these results also hold on the larger domain \mathcal{A} , that is, when externalities are present. However, before we can even start to answer these questions, we have to define stable outcomes in these more general problems. Example 1 shows that this is not straightforward.

⁴ That $n = m$ is without loss as we otherwise can balance the count with agents that create no value in any pair.

Table 1 Values for Example 1

Matching	Pair 1	Pair 2	Pair 3
$\mu_1 = \{(f_1, w_1), (f_2, w_2), (f_3, w_3)\}$	2	2	1
$\mu_2 = \{(f_1, w_2), (f_2, w_3), (f_3, w_1)\}$	2	0	2
$\mu_3 = \{(f_1, w_1), (f_2, w_3), (f_3, w_2)\}$	2	2	1

For instance, the “0” indicates that f_2 and w_3 are matched at μ_2 and generate a value of $\alpha_{f_2 w_3}(\mu_2) = 0$

Example 1 (What outcomes are stable?) Let $F = \{f_1, f_2, f_3\}$ and $W = \{w_1, w_2, w_3\}$. Table 1 displays the values created by the different pairs; all other values are zero. Is the outcome (μ_1, u, v) with $u = (1, 1, 1)$ and $v = (1, 1, 0)$ stable?

Except for f_2 and w_3 , no pair has anything to gain from deviating from (μ_1, u, v) . The stability of (μ_1, u, v) , therefore, boils down to whether f_2 and w_3 object to it. We have $u_{f_2} + v_{w_3} = 1 + 0$ and

$$\alpha_{f_2 w_3}(\mu_2) = 0 < 1 < 2 = \alpha_{f_2 w_3}(\mu_3).$$

Therefore, it makes sense for f_2 and w_3 to break up their current partnerships and match with one another if the matching formed thereupon is μ_3 , but not if it is μ_2 . However, this is out of their control. Hence, whether f_2 and w_3 block the outcome depends on whether they are more optimistic (expect μ_3 to be formed) than pessimistic (expect μ_2). \circ

We wish to formalize the insights of Example 1. Agents have expectations on what will occur as a consequence of them blocking a matching. A pair of agents **block** an outcome if their “blocking threshold” exceeds their joint payoffs. For each $(f, w) \in F \times W$, define the optimistic value $\omega_{fw} \in \mathbb{R}$ as the largest value that the pair can achieve. Conversely, define the pessimistic value $\pi_{fw} \in \mathbb{R}$ as the smallest value that the pair can achieve (that is, the value that the pair can guarantee itself). In Example 1, $\omega_{f_2 w_3} = \alpha_{f_2 w_3}(\mu_3) = 2$ and $\pi_{f_2 w_3} = \alpha_{f_2 w_3}(\mu_2) = 0$. Formally,

$$\begin{aligned}\omega_{fw} &= \max_{\mu \in \mathcal{M}_{fw}} \alpha_{fw}(\mu) \\ \pi_{fw} &= \min_{\mu \in \mathcal{M}_{fw}} \alpha_{fw}(\mu).\end{aligned}$$

Certainly, f and w have strong reasons to object to an outcome for which $u_f + v_w < \pi_{fw}$ —if they break up their current partnerships to instead form one together, they are better off no matter how the other agents react. For each $(f, w) \in F \times W$, $\lambda_{fw} : \mathcal{M}_{-fw} \rightarrow [0, 1]$ is used to determine the **blocking threshold** $b_{fw}(\mu) \in \mathbb{R}$.⁵

$$b_{fw}(\mu) = \lambda_{fw}(\mu) \cdot \omega_{fw} + (1 - \lambda_{fw}(\mu)) \cdot \pi_{fw}.$$

⁵ We may assume that f and w have different thresholds. However, as f and w block only if they agree on doing so, it is only the smaller of the individual blocking thresholds that is relevant.

A pair $(f, w) \in F \times W$ is **pessimistic** at $\mu \in \mathcal{M}_{-fw}$ if $\lambda_{fw}(\mu) = 0$. A **game with externalities** is (A, λ) , where $A \in \mathcal{A}$ and $\lambda \equiv (\lambda_{fw}(\mu))_{f \in F, w \in W, \mu \in \mathcal{M}_{-fw}}$. The **collection of games with externalities** is \mathcal{E} . A pair $(f, w) \in F \times W$ **blocks** the outcome $(\mu, u, v) \in \mathcal{M} \times \mathbb{R}_+^m \times \mathbb{R}_+^n$ of $(A, \lambda) \in \mathcal{E}$ if $u_f + v_w < b_{fw}(\mu)$. An outcome (μ, u, v) is **stable** in (A, λ) if no pair blocks it.

3 Results

We first observe that the set of stable outcomes for a game with externalities is related to that of a particular game without externalities.

Observation 1 *An outcome $(\mu, u, v) \in \mathcal{M} \times \mathbb{R}_+^m \times \mathbb{R}_+^n$ is stable in $(A, \lambda) \in \mathcal{E}$ if and only if (μ, u, v) is stable in $B^\mu \in \mathcal{A}^0$, where $B^\mu = (\beta_{fw}^\mu)_{f \in F, w \in W}$ and*

$$\beta_{fw}^\mu = \begin{cases} \alpha_{fw}(\mu) & \text{if } (f, w) \in \mu \\ b_{fw}(\mu) & \text{if } (f, w) \notin \mu. \end{cases}$$

Proof Let (μ, u, v) be stable in (A, λ) . We proceed in two steps. First, we show that (μ, u, v) is an outcome of B^μ . Then we show that (μ, u, v) is stable in B^μ .

As (μ, u, v) is an outcome of (A, λ) , for each $f \in F$ and $w = \mu(f)$, $u_f + v_w = \alpha_{fw}(\mu)$. By construction, $\alpha_{fw}(\mu) = \beta_{fw}^\mu$. Therefore, $u_f + v_w = \beta_{fw}^\mu$, so (μ, u, v) is an outcome of B^μ .

As (μ, u, v) is stable in (A, λ) , for each $(f, w) \in F \times W$ such that $w \neq \mu(f)$, $u_f + v_w \geq b_{fw}(\mu)$. By construction, $b_{fw}(\mu) = \beta_{fw}^\mu$. Then $u_f + v_w \geq \beta_{fw}^\mu$, so (μ, u, v) is stable in B^μ .

For the other direction, suppose that (μ, u, v) is not stable in (A, λ) . Then there is $(f, w) \in F \times W$ such that $w \neq \mu(f)$ and $u_f + v_w < b_{fw}(\mu) = \beta_{fw}^\mu$. But then f and w block (μ, u, v) in B^μ . \square

This result *does not* imply that there always are stable outcomes in games with externalities. Surely, B^μ has stable outcomes as it has no externalities (Proposition 1). However, if these outcomes use matchings other than μ , then none of them needs to be stable in (A, λ) .

3.1 Existence

Next, we state a sufficient condition for existence of stable outcomes. Namely, there are stable outcomes if agents are pessimistic and hence conservative in forming blocking pairs.

Proposition 2 (Sasaki and Toda 1996, Theorem 6.1) *Fix λ such that pairs always are pessimistic: for each $(f, w) \in F \times W$ and $\mu \in \mathcal{M}_{-fw}$, $\lambda_{fw}(\mu) = 0$. Then, for each $A \in \mathcal{A}$, there is a stable outcome in (A, λ) .*

Our main result shows that even the “slightest” optimism can deter the existence of stable outcomes. This result is strong: it requires just *one* pair to be non-pessimistic at just *one* matching.

Theorem 1 Fix λ such that pairs are not always pessimistic: there is $(f, w) \in F \times W$, $\mu \in \mathcal{M}_{-fw}$, and $\varepsilon > 0$ such that $\lambda_{fw}(\mu) = \varepsilon$. Then there is $A \in \mathcal{A}$ such that (A, λ) has no stable outcome.

Proof Let $F = \{f_1, f_2, \dots, f_n\}$ and $W = \{w_1, w_2, \dots, w_n\}$. Assume $(f_1, w_1) \in F \times W$ is not pessimistic at $\hat{\mu} \in \mathcal{M}_{-f_1w_1}$, say $\lambda_{f_1w_1}(\hat{\mu}) = \varepsilon$. Without loss, define $\hat{\mu}$ and $\tilde{\mu} \in \mathcal{M}$ as follows:

$$\begin{aligned}\hat{\mu} &= \{(f_1, w_2), (f_2, w_3), \dots, (f_{n-1}, w_n), (f_n, w_1)\} \\ \tilde{\mu} &= \{(f_1, w_1), (f_2, w_2), \dots, (f_n, w_n)\}\end{aligned}$$

Define $A \in \mathcal{A}$ such that, for each $f_i \in F$ and $\mu \in \mathcal{M}_{f_iw_{i+1}}$, $\alpha_{f_iw_{i+1}}(\mu) = 1 \pmod{n}$. Moreover, $\alpha_{f_1w_1}(\tilde{\mu}) = 2/\varepsilon + 1$, and let all other values be 0. Then

$$\begin{aligned}b_{f_1w_1}(\hat{\mu}) &= [\lambda_{f_1w_1}(\hat{\mu})] \cdot \omega_{f_1w_1} + [(1 - \lambda_{f_1w_1}(\hat{\mu}))] \cdot \pi_{f_1w_1} \\ &= \varepsilon \cdot (2/\varepsilon + 1) + (1 - \varepsilon) \cdot 0 = 2 + \varepsilon.\end{aligned}$$

Additionally, for each $f_i \in F$ and $\mu \in \mathcal{M}_{-f_iw_{i+1}}$, $\omega_{f_iw_{i+1}} = \pi_{f_iw_{i+1}} = 1$, so $b_{f_iw_{i+1}}(\mu) = 1$.

Let $(u, v) \in \mathbb{R}_+^m \times \mathbb{R}_+^n$. Consider $(\hat{\mu}, u, v)$. As $\hat{\mu}(f_1) = w_2$, $u_{f_1} + v_{w_2} = \alpha_{f_1w_2}(\hat{\mu}) = 1$. As $\hat{\mu}(f_n) = w_1$, $u_{f_n} + v_{w_1} = \alpha_{f_nw_1}(\hat{\mu}) = 1$. As $u_{f_n} \geq 0$ and $v_{w_n} \geq 0$, $u_{f_1} + v_{w_1} \leq 2 < b_{f_1w_1}(\hat{\mu})$. Hence, f_1 and w_1 block $(\hat{\mu}, u, v)$.

Consider (μ, u, v) for $\mu \neq \hat{\mu}$. Then there is $f_i \in F$ such that $\mu(f_i) \neq w_{i+1}$. Without loss, let $\mu(f_i) = w_j$ and $\mu(f_k) = w_{i+1}$. As $\mu(f_i) = w_j$, $u_{f_i} + v_{w_j} = \alpha_{f_iw_j}(\mu) = 0$. As $\mu(f_k) = w_{i+1}$, $u_{f_k} + v_{w_{i+1}} = \alpha_{f_kw_{i+1}}(\mu) = 0$. Then $u_{f_i} = u_{f_k} = v_{w_j} = v_{w_{i+1}} = 0$, and therefore, $u_{f_i} + v_{w_{i+1}} = 0 < b_{f_iw_{i+1}}(\mu)$. Hence, f_i and w_{i+1} block (μ, u, v) .

As this exhausts all possibilities, there is no stable outcome. \square

For Theorem 1, we exploit that externalities may be arbitrarily large. In Example 2, we show that, if agents are sufficiently optimistic, then there need not exist stable outcomes even if externalities are vanishingly small.

Example 2 (Minimal externalities, no stable outcome) Consider a game with externalities where pairs are “sufficiently” optimistic. For each $(f, w) \in F \times W$ and $\mu \in \mathcal{M}_{-fw}$, $\lambda_{fw}(\mu) > (n - 1)/n$. Values are symmetric and externalities are vanishingly small: for some $\varepsilon > 0$ and each $(f, w) \in F \times W$,

$$\{\alpha_{fw}(\mu) : \mu \in \mathcal{M}_{fw}\} = \{1, 1 + \varepsilon\}.$$

Then, for each $(f, w) \in F \times W$ and $\mu \in \mathcal{M}_{-fw}$, $\omega_{fw} = 1 + \varepsilon$. Assume, for each $\mu \in \mathcal{M}$, that there is $(f, w) \in \mu$ such that $\alpha_{fw}(\mu) = 1$. Then there is no stable outcome.⁶ The externalities vanish both in relative and absolute terms:

⁶ The smallest value needed to satisfy all pairs exceeds $n[(n - 1)/n] \cdot (1 + \varepsilon) + [1/n] \cdot 1$, which simplifies to $n(1 + \varepsilon) - \varepsilon$. But this is also an upper bound on the total value available.

$$\frac{\omega_{fw}}{\pi_{fw}} = \frac{1 + \varepsilon}{1} \rightarrow 1 \quad \text{as } \varepsilon \rightarrow 0 \quad \omega_{fw} - \pi_{fw} = (1 + \varepsilon) - 1 \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0.$$

Hence, even the smallest of externalities can be problematic if agents are optimistic. \circ

3.2 Efficiency

In contrast to games without externalities, efficiency and stability now no longer go hand in hand. Sasaki and Toda (1996, Example 6.1) show that inefficient matchings may be stable. To complement this finding, Example 3 shows that efficient matchings may be unstable.

Example 3 (Unstable efficient matching) Let $F = \{f_1, f_2, f_3\}$, $W = \{w_1, w_2, w_3\}$, and values $\alpha_{fw}(\mu) = 0$ except the following:

$$\begin{aligned} \alpha_{f_1 w_1}(\mu_1) &= 2 \quad \text{for } \mu_1 = \{(f_1, w_1), (f_2, w_2), (f_3, w_3)\} \\ \alpha_{f_2 w_3}(\mu_2) &= 1 \quad \text{for } \mu_2 = \{(f_1, w_1), (f_2, w_3), (f_3, w_2)\} \\ \alpha_{f_2 w_3}(\mu_3) &= 1 \quad \text{for } \mu_3 = \{(f_1, w_2), (f_2, w_3), (f_3, w_1)\}. \end{aligned}$$

Then μ_1 is efficient though agents f_2 and w_3 block each outcome (μ_1, u, v) . The inefficient outcome (μ_2, u, v) is stable for $u = (0, 1, 0)$ and $v = (0, 0, 0)$. \circ

3.3 The structure of the set of stable outcomes

Without externalities the set of stable outcomes forms a complete lattice with respect to the agents' payoffs. It has two extreme points: one which all firms prefer to all other stable outcomes, one which all workers prefer to all other stable outcomes. By the intuition of Observation 1, we retrieve an immediate corollary of this result if we restrict attention to *specific* matchings: for each $(A, \lambda) \in \mathcal{E}$ and $\mu \in \mathcal{M}$,

$$\{(u, v) \in \mathbb{R}_+^m \times \mathbb{R}_+^n : (\mu, u, v) \text{ is stable in } (A, \lambda)\}$$

forms a complete lattice. Thus, if (A, λ) has a unique stable matching, then the game's set of stable outcomes forms a lattice. However, Example 4 shows that, if there are multiple stable matchings, then the set of stable outcomes may lose this structure.

Example 4 (Set of stable outcomes, no lattice structure) Let $F = \{f_1, f_2, f_3\}$ and $W = \{w_1, w_2, w_3\}$. All pairs are pessimistic; values are in Table 2.

There are two stable matchings: μ_1 and μ_4 . The set of stable payoffs compatible with μ_1 forms a lattice with (firm-) minimal element $u = (1, 0, 0)$ and $v = (7, 6, 4)$. This set is disjoint from the set of stable outcomes compatible with μ_4 . This latter set has minimal element $u' = (0, 0, 1)$ with $v' = (5, 4, 8)$. For neither matching, the *meet* $u \wedge u' = (0, 0, 0)$ is stable. See Fig. 1 for an illustration.

As agents get more optimistic, they object to more outcomes, so the set of stable outcomes shrinks. For concreteness, suppose for each $(f, w) \in F \times W$ and $\mu \in \mathcal{M}_{-fw}$, $\lambda_{fw}(\mu) = 0.4$. Then only μ_4 is stable. \circ

Table 2 Values for Example 4

Matching	Pair 1	Pair 2	Pair 3
$\mu_1 = \{(f_1, w_1), (f_2, w_2), (f_3, w_3)\}$	8	6	4
$\mu_2 = \{(f_1, w_1), (f_2, w_3), (f_3, w_2)\}$	3	4	5
$\mu_3 = \{(f_1, w_2), (f_2, w_1), (f_3, w_3)\}$	6	4	4
$\mu_4 = \{(f_1, w_2), (f_2, w_3), (f_3, w_1)\}$	4	8	6
$\mu_5 = \{(f_1, w_3), (f_2, w_1), (f_3, w_2)\}$	5	2	5
$\mu_6 = \{(f_1, w_3), (f_2, w_2), (f_3, w_1)\}$	5	2	5

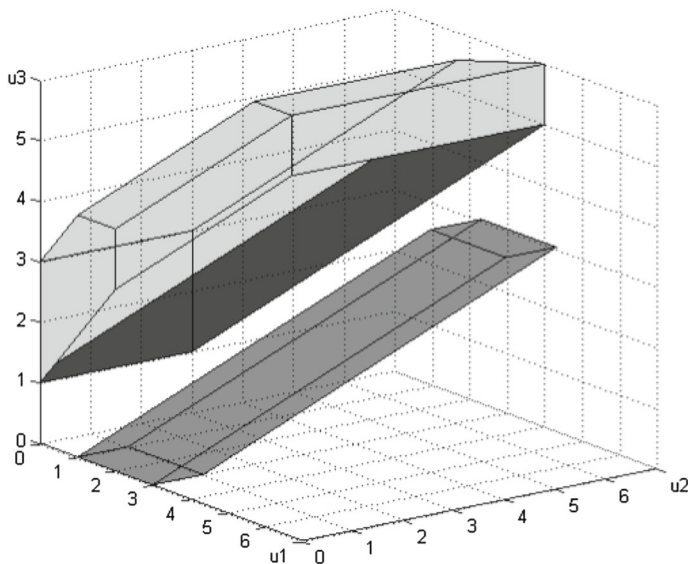


Fig. 1 (Set of stable payoffs for Example 4.) The dark grey area shows the lattice structure of the stable payoffs compatible with μ_1 when agents are pessimistic. The light grey area is the set of stable payoffs compatible with μ_4 . The black area is the set of stable payoffs when agents are less pessimistic, namely, for each $(f, w) \in F \times W$ and $\mu \in \mathcal{M}_{-fw}$, $\lambda_{fw}(\mu) = 0.4$. Importantly, the two grey areas are disjoint

Without externalities, we can “swap” stable matchings: if the outcomes (μ, u, v) and (μ', u', v') are stable in $A \in \mathcal{A}^0$, then so is (μ', u, v) . It follows immediately from Example 4 that this does not extend to games with externalities.

4 Imposing more structure

In the proof of Theorem 1 and in Example 2, we exploit either (i) large externalities when agents are optimistic or (ii) (almost) optimistic agents when externalities are small. In this section, we impose more structure on the values and the externalities and ask whether this overturns the negative findings. The setting is highly restrictive on

Table 3 Values for Example 5

Matching	Pair 1	Pair 2	Pair 3
$\mu_1 = \{(f_1, w_1), (f_2, w_2), (f_3, w_3)\}$	1	8	7
$\mu_2 = \{(f_1, w_1), (f_2, w_3), (f_3, w_2)\}$	3	9	6
$\mu_3 = \{(f_1, w_2), (f_2, w_1), (f_3, w_3)\}$	2	7	5
$\mu_4 = \{(f_1, w_2), (f_2, w_3), (f_3, w_1)\}$	6	9	3
$\mu_5 = \{(f_1, w_3), (f_2, w_1), (f_3, w_2)\}$	5	7	2
$\mu_6 = \{(f_1, w_3), (f_2, w_2), (f_3, w_1)\}$	7	8	1

purpose: we aim to show that the negative results persist even if values are constructed “realistically”.

The skill of worker w is $s(w) \in \mathbb{R}$: no matter where w is employed, she adds value $s(w)$. Firms produce different products, some complement each other, some substitute. The “degree of complementarity” is captured by $\kappa : F \times F \rightarrow \mathbb{R}$. If firms f and f' produce complementary goods, then $\kappa(f, f') > 0$. If they produce substitutes, $\kappa(f, f') < 0$. For each $f \in F$, $\kappa(f, f) = 1$. The value created by $(f, w) \in F \times W$ at $\mu \in \mathcal{M}_{fw}$ is

$$\alpha_{fw}(\mu) = \sum_{f' \in F} \kappa(f, f') \cdot s(\mu(f')).$$

Example 5 (Negative results not overturned by more structure) Let $F = \{f_1, f_2, f_3\}$ and $W = \{w_1, w_2, w_3\}$. For simplicity, the skill of $w_i \in W$ is $s(w_i) = 2i$. That is, w_1 is the least productive worker, w_3 the most. Firms f_1 and f_3 produce substitutable goods (different brands of phones), say $\kappa(f_1, f_3) = -0.5$. Firm f_2 produces a complementary good (phone accessories), say $\kappa(f_1, f_2) = \kappa(f_2, f_3) = 0.5$.⁷ Therefore, at $\mu_1 = \{(f_1, w_1), (f_2, w_2), (f_3, w_3)\}$, (f_1, w_1) creates a value of 1:

$$\begin{aligned} \alpha_{f_1 w_1}(\mu_1) &= \kappa(f_1, f_1) \cdot s(\mu_1(f_1)) + \kappa(f_1, f_2) \cdot s(\mu_1(f_2)) + \kappa(f_1, f_3) \cdot s(\mu_1(f_3)) \\ &= \underbrace{1 \cdot s(w_1)}_2 + \underbrace{1/2 \cdot s(w_2)}_2 + \underbrace{(-1/2) \cdot s(w_3)}_{-3} \\ &= 1. \end{aligned}$$

All values are displayed in Table 3.

If pairs always are pessimistic, then each matching is stable. However, when no pair is pessimistic, that is, when there is $\varepsilon > 0$ such that, for each $(f, w) \in F \times W$ and $\mu \in \mathcal{M}_{fw}$, $\lambda_{fw}(\mu) \geq \varepsilon$, then there is no stable outcome. We only give proof to this claim for μ_1 . The exercise can be repeated for the other matchings.

⁷ Number are chosen to simplify the example; more “varied” example are available upon request.

Suppose, to obtain a contradiction, that (μ_1, u, v) is stable. For f_1 and w_2 not to block,

$$u_{f_1} + v_{w_2} = u_{f_1} + (8 - u_{f_2}) \geq b_{f_1 w_2}(\mu_1) \geq \varepsilon \cdot \omega_{f_1 w_2} + (1 - \varepsilon) \cdot \pi_{f_1 w_2} > \pi_{f_1 w_2} = 2.$$

Then $u_{f_2} - u_{f_1} < 6$. For f_2 and w_1 not to block,

$$u_{f_2} + v_{w_1} = u_{f_2} + (1 - u_{f_1}) \geq b_{f_2 w_1}(\mu_1) = 7 \Leftrightarrow u_{f_2} - u_{f_1} \geq 6.$$

Thus, no matter the outcome (μ_1, u, v) , either f_1 and w_2 or f_2 and w_1 block. \circ

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aumann, R.J., Peleg, B.: Von Neumann–Morgenstern solutions to cooperative games without side payments. *Bull. Am. Math. Soc.* **66**, 173–179 (1960)
- Chen, B.: Assignment games with externalities and matching-based Cournot competition. *Bonn Econ Discussion Papers* 08/2013 (2013)
- Dutta, B., Massó, J.: Stability of matchings when individuals have preferences over colleagues. *J. Econ. Theory* **75**(2), 464–475 (1997)
- Echenique, F., Yenmez, M.B.: A solution to matching with preferences over colleagues. *Games Econ. Behav.* **59**(1), 46–71 (2007)
- Eriksson, K., Jansson, F., Vetander, T.: The assignment game with negative externalities and bounded rationality. *Int. Game Theory Rev.* **13**(04), 443–459 (2011)
- Funaki, Y., Yamato, T.: The core of an economy with a common pool resource: a partition function form approach. *Int. J. Game Theory* **28**(2), 157–171. <http://ideas.repec.org/a/spr/jogath/v28y1999i2p157-171.html> (1999)
- Gale, D., Shapley, L.S.: College admissions and the stability of marriage. *Am. Math. Mon.* **69**(1), 9–15 (1962)
- Hafalir, I.E.: Stability of marriage with externalities. *Int. J. Game Theory* **37**(3), 353–369 (2008)
- Kóczy, L.: A recursive core for partition function form games. *Theory Decis.* **63**(1), 41–51 (2007)
- Koopmans, T., Beckmann, M.: Assignment problems and the location of economic activities. *Econometrica* **25**, 53–76 (1957)
- Li, S.: Competitive matching equilibrium and multiple principal-agent models. Technical report, Center for Economic Research, Department of Economics, University of Minnesota (1993)
- Mumcu, A., Saglam, I.: The core of a housing market with externalities. *Econ. Bull.* **3**(57), 1–5. <http://ideas.repec.org/a/ebull/ecbull/eb-07c70026.html> (2007)
- Pycia, M., Yenmez, M.B.: Matching with Externalities. Mimeo, New York (2016)
- Roth, A.E., Sotomayor, M.A.O.: Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis, vol. 18. Cambridge University Press, Cambridge (1992)
- Sasaki, H., Toda, M.: Two-sided matching problems with externalities. *J. Econ. Theory* **70**(1), 93–108 (1996)
- Shapley, L.S., Shubik, M.: The assignment game I: the core. *Int. J. Game Theory* **1**(1), 111–130 (1971)
- Shenoy, P.P.: A dynamic solution concept for abstract games. *J. Optim. Theory Appl.* **32**(2), 151–169 (1980)